



## RAW SEQUENCE LISTING

DATE: 01/17/2003

PATENT APPLICATION: PCT/US02/40887

TIME: 15:17:43

Input Set : A:\257618.txt

Output Set: N:\CRF4\01172003\PU40887.raw

3 <110> APPLICANT: Boyce Thompson Institute for Plant Research  
 4 Pioneer Hi-Bred International, Inc.  
 5 May, Gregory D  
 6 Baszczynski, Christopher L  
 7 Zhu, Tong  
 8 Kipp, Peter B  
 9 Mahajan, Pramod B  
 11 <120> TITLE OF INVENTION: PLANT MSH2 SEQUENCES AND METHODS OF USE  
 13 <130> FILE REFERENCE: 42960/257618  
 C--> 15 <140> CURRENT APPLICATION NUMBER: PCT/US02/40887  
 C--> 15 <141> CURRENT FILING DATE: 2002-12-20  
 15 <150> PRIOR APPLICATION NUMBER: US 10/029,065  
 16 <151> PRIOR FILING DATE: 2001-12-20  
 18 <160> NUMBER OF SEQ ID NOS: 42  
 20 <170> SOFTWARE: PatentIn version 3.1  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 3033  
 24 <212> TYPE: DNA  
 25 <213> ORGANISM: Nicotiana tabacum  
 27 <220> FEATURE:  
 28 <221> NAME/KEY: CDS  
 29 <222> LOCATION: (22)..(2838)  
 30 <223> OTHER INFORMATION:  
 W--> 32 <400> 1  
 33 ataaagggtta aagaaaaaaa a atg aat gaa aat ttg gag gaa cag agc aag 51  
 34 Met Asn Glu Asn Leu Glu Glu Gln Ser Lys  
 35 1 5 10  
 37 ctt ccc gag ctt aaa ctg gat gct aag caa gct caa gga ttt ctc tca 99  
 38 Leu Pro Glu Leu Lys Leu Asp Ala Lys Gln Ala Gln Gly Phe Leu Ser  
 39 15 20 25  
 41 ttc ttc aaa acc ctg ccc aag gac cct agg gca gtt cgc ctc ttt gat 147  
 42 Phe Phe Lys Thr Leu Pro Lys Asp Pro Arg Ala Val Arg Leu Phe Asp  
 43 30 35 40  
 45 cgt cgg gac tat tat aca tct cat gga gat gat gca act ttc att gca 195  
 46 Arg Arg Asp Tyr Tyr Thr Ser His Gly Asp Asp Ala Thr Phe Ile Ala  
 47 45 50 55  
 49 gag aca tat tac cac aca aca act gcg tta cga cag ttg ggt aat aga 243  
 50 Glu Thr Tyr Tyr His Thr Thr Thr Ala Leu Arg Gln Leu Gly Asn Arg  
 51 60 65 70  
 53 gct gat gcc ctt tcc agt gtt agt gtg agt aga aac atg ttt gaa aca 291  
 54 Ala Asp Ala Leu Ser Ser Val Ser Val Ser Arg Asn Met Phe Glu Thr  
 55 75 80 85 90  
 57 ata gct cgt gac att ctc ttg gag aga atg gac cgt act ctt gaa cta 339

ENTERED

## RAW SEQUENCE LISTING

DATE: 01/17/2003

PATENT APPLICATION: PCT/US02/40887

TIME: 15:17:43

Input Set : A:\257618.txt

Output Set: N:\CRF4\01172003\PU40887.raw

58	Ile	Ala	Arg	Asp	Ile	Leu	Leu	Glu	Arg	Met	Asp	Arg	Thr	Leu	Glu	Leu	
59					95					100					105		
61	tat	gag	ggc	agt	ggt	tca	aac	tgg	aga	ctg	gta	aaa	agt	gga	acc	cca	387
62	Tyr	Glu	Gly	Ser	Gly	Ser	Asn	Trp	Arg	Leu	Val	Lys	Ser	Gly	Thr	Pro	
63				110					115					120			
65	ggg	aat	ctt	gga	agt	ttt	gag	gat	att	ctg	ttt	gct	aat	aat	gaa	atg	435
66	Gly	Asn	Leu	Gly	Ser	Phe	Glu	Asp	Ile	Leu	Phe	Ala	Asn	Asn	Glu	Met	
67			125					130					135				
69	caa	aat	tct	ccg	gtg	att	gct	gct	ctt	gct	cca	aac	ttc	ggt	cag	aat	483
70	Gln	Asn	Ser	Pro	Val	Ile	Ala	Ala	Leu	Ala	Pro	Asn	Phe	Gly	Gln	Asn	
71		140					145					150					
73	gga	tgt	gaa	gtt	ggc	tta	ggc	tat	gtt	gat	att	act	aag	aga	gtc	ctt	531
74	Gly	Cys	Glu	Val	Gly	Leu	Gly	Tyr	Val	Asp	Ile	Thr	Lys	Arg	Val	Leu	
75	155					160				165					170		
77	ggt	tta	aca	gaa	ttt	cta	gat	gat	agc	cac	ttc	aca	aat	ttg	gag	tct	579
78	Gly	Leu	Thr	Glu	Phe	Leu	Asp	Asp	Ser	His	Phe	Thr	Asn	Leu	Glu	Ser	
79				175					180					185			
81	gct	ttg	gtt	gct	ctt	ggt	tgc	aga	gaa	tgt	ctt	gta	cca	gcg	gag	act	627
82	Ala	Leu	Val	Ala	Leu	Gly	Cys	Arg	Glu	Cys	Leu	Val	Pro	Ala	Glu	Thr	
83			190					195					200				
85	ggc	aaa	tcc	agt	gaa	tac	agg	cct	atg	ttt	gat	gca	ata	tct	aga	tgc	675
86	Gly	Lys	Ser	Ser	Glu	Tyr	Arg	Pro	Met	Phe	Asp	Ala	Ile	Ser	Arg	Cys	
87		205					210					215					
89	ggc	gtg	atg	gta	act	gaa	aga	aag	aaa	act	gaa	ttt	aaa	ggg	aga	gat	723
90	Gly	Val	Met	Val	Thr	Glu	Arg	Lys	Lys	Thr	Glu	Phe	Lys	Gly	Arg	Asp	
91		220					225				230						
93	ttg	gta	cag	gat	ctt	ggt	agg	ctc	gtc	aag	ggt	tca	gta	gaa	cct	gtt	771
94	Leu	Val	Gln	Asp	Leu	Gly	Arg	Leu	Val	Lys	Gly	Ser	Val	Glu	Pro	Val	
95	235				240					245				250			
97	cga	gat	ttg	gtc	tct	ggg	ttc	gaa	tgt	gca	tca	ggc	gct	ttg	ggg	tgc	819
98	Arg	Asp	Leu	Val	Ser	Gly	Phe	Glu	Cys	Ala	Ser	Gly	Ala	Leu	Gly	Cys	
99			255					260				265					
101	ata	ctt	tct	tat	gca	gaa	cta	ctt	gcg	gat	gag	agc	aac	tat	gga	aac	867
102	Ile	Leu	Ser	Tyr	Ala	Glu	Leu	Leu	Ala	Asp	Glu	Ser	Asn	Tyr	Gly	Asn	
103			270					275					280				
105	tat	aca	gtc	aaa	caa	tac	aac	ctc	aat	agt	tac	atg	aga	tta	gat	tct	915
106	Tyr	Thr	Val	Lys	Gln	Tyr	Asn	Leu	Asn	Ser	Tyr	Met	Arg	Leu	Asp	Ser	
107		285					290					295					
109	gct	gct	atg	aga	gca	ctg	aat	gtt	atg	gag	agc	aaa	tca	gat	gct	aat	963
110	Ala	Ala	Met	Arg	Ala	Leu	Asn	Val	Met	Glu	Ser	Lys	Ser	Asp	Ala	Asn	
111		300				305						310					
113	aaa	aat	ttt	agc	ttg	ttc	ggt	ctg	atg	aat	aga	acg	tgt	act	gct	gga	1011
114	Lys	Asn	Phe	Ser	Leu	Phe	Gly	Leu	Met	Asn	Arg	Thr	Cys	Thr	Ala	Gly	
115	315				320					325				330			
117	atg	ggt	aaa	agg	tta	ttg	cac	atg	tgg	ctg	aag	caa	cct	tta	cta	gat	1059
118	Met	Gly	Lys	Arg	Leu	Leu	His	Met	Trp	Leu	Lys	Gln	Pro	Leu	Leu	Asp	
119			335					340					345				
121	gta	gaa	gag	att	aac	tgt	agg	ctg	gat	tta	gtt	caa	tca	ttc	gtg	gag	1107
122	Val	Glu	Glu	Ile	Asn	Cys	Arg	Leu	Asp	Leu	Val	Gln	Ser	Phe	Val	Glu	



## RAW SEQUENCE LISTING

PATENT APPLICATION: PCT/US02/40887

DATE: 01/17/2003

TIME: 15:17:43

Input Set : A:\257618.txt

Output Set: N:\CRF4\01172003\PU40887.raw

189	cca	aat	atc	agt	cca	cca	gat	aca	gga	gat	att	ata	ctt	gaa	ggg	tgt	1923
190	Pro	Asn	Ile	Ser	Pro	Pro	Asp	Thr	Gly	Asp	Ile	Ile	Leu	Glu	Gly	Cys	
191	620						625					630					
193	agg	cat	cct	tgt	gtg	gaa	gct	caa	gat	tgg	gtt	aac	tcc	att	cct	aat	1971
194	Arg	His	Pro	Cys	Val	Glu	Ala	Gln	Asp	Trp	Val	Asn	Ser	Ile	Pro	Asn	
195	635					640					645					650	
197	gac	tgt	aga	cta	gtt	agg	gga	gag	agt	tgg	ttt	cag	att	atc	aca	ggc	2019
198	Asp	Cys	Arg	Leu	Val	Arg	Gly	Glu	Ser	Trp	Phe	Gln	Ile	Ile	Thr	Gly	
199					655					660						665	
201	cct	aac	atg	ggt	gga	aag	tcg	acc	tac	att	cgg	cag	gtt	ggt	gtg	aat	2067
202	Pro	Asn	Met	Gly	Gly	Lys	Ser	Thr	Tyr	Ile	Arg	Gln	Val	Gly	Val	Asn	
203				670					675					680			
205	gtc	ctg	atg	gcc	caa	gtt	ggc	tcg	ttt	gtt	cca	tgt	gac	aat	gct	acc	2115
206	Val	Leu	Met	Ala	Gln	Val	Gly	Ser	Phe	Val	Pro	Cys	Asp	Asn	Ala	Thr	
207			685					690					695				
209	att	tct	att	cgt	gat	tgt	att	ttt	gct	cgt	gtt	ggc	gct	gga	gat	tgc	2163
210	Ile	Ser	Ile	Arg	Asp	Cys	Ile	Phe	Ala	Arg	Val	Gly	Ala	Gly	Asp	Cys	
211		700					705					710					
213	cag	ctg	aga	gga	gtt	tct	act	ttt	atg	caa	gag	atg	ctt	gag	act	gca	2211
214	Gln	Leu	Arg	Gly	Val	Ser	Thr	Phe	Met	Gln	Glu	Met	Leu	Glu	Thr	Ala	
215	715					720					725					730	
217	tcg	atc	ttg	aaa	gga	gct	act	gat	aga	tca	ttg	att	ata	att	gat	gag	2259
218	Ser	Ile	Leu	Lys	Gly	Ala	Thr	Asp	Arg	Ser	Leu	Ile	Ile	Ile	Asp	Glu	
219				735						740						745	
221	ttg	ggc	cgt	ggg	aca	tca	acc	tac	gat	ggc	ttt	ggt	tta	gct	tgg	gct	2307
222	Leu	Gly	Arg	Gly	Thr	Ser	Thr	Tyr	Asp	Gly	Phe	Gly	Leu	Ala	Trp	Ala	
223			750						755					760			
225	att	tgt	gag	cac	att	gtt	gaa	gaa	att	aaa	gca	cca	aca	ttg	ttt	gcc	2355
226	Ile	Cys	Glu	His	Ile	Val	Glu	Glu	Ile	Lys	Ala	Pro	Thr	Leu	Phe	Ala	
227		765						770					775				
229	act	cac	ttt	cat	gag	ctg	act	gca	tta	gcc	aac	aag	aat	gga	gac	aat	2403
230	Thr	His	Phe	His	Glu	Leu	Thr	Ala	Leu	Ala	Asn	Lys	Asn	Gly	Asp	Asn	
231		780					785					790					
233	gga	cat	aag	aaa	aat	gct	ggg	ata	gca	aat	ttt	cat	gtt	ttt	gca	cac	2451
234	Gly	His	Lys	Lys	Asn	Ala	Gly	Ile	Ala	Asn	Phe	His	Val	Phe	Ala	His	
235	795					800					805					810	
237	att	gac	cct	tct	aat	cgc	aag	cta	act	atg	ctt	tac	aag	gtt	cac	cca	2499
238	Ile	Asp	Pro	Ser	Asn	Arg	Lys	Leu	Thr	Met	Leu	Tyr	Lys	Val	His	Pro	
239				815						820				825			
241	ggt	gct	tgt	gat	cag	agt	ttt	ggt	att	cat	gtt	gct	gaa	ttt	gca	aat	2547
242	Gly	Ala	Cys	Asp	Gln	Ser	Phe	Gly	Ile	His	Val	Ala	Glu	Phe	Ala	Asn	
243			830						835					840			
245	ttt	cca	ccg	agt	gtt	gtg	gct	ctg	gct	aga	gaa	aag	gca	tct	gag	ttg	2595
246	Phe	Pro	Pro	Ser	Val	Val	Ala	Leu	Ala	Arg	Glu	Lys	Ala	Ser	Glu	Leu	
247		845						850					855				
249	gag	gat	ttc	tct	cct	att	gcc	ata	att	cca	aat	gac	att	aaa	gag	gca	2643
250	Glu	Asp	Phe	Ser	Pro	Ile	Ala	Ile	Ile	Pro	Asn	Asp	Ile	Lys	Glu	Ala	
251		860					865					870					
253	gct	tca	aaa	cgg	aag	aga	gaa	ttt	gac	cgc	cat	gac	gtg	tct	aga	ggt	2691

## RAW SEQUENCE LISTING

PATENT APPLICATION: PCT/US02/40887

DATE: 01/17/2003

TIME: 15:17:43

Input Set : A:\257618.txt

Output Set: N:\CRF4\01172003\PU40887.raw

```

254 Ala Ser Lys Arg Lys Arg Glu Phe Asp Arg His Asp Val Ser Arg Gly
255 875                               880                               885                               890
257 act gcc aga gct cgg caa ttc tta cag gat ttc gct cag ttg cca ctg      2739
258 Thr Ala Arg Ala Arg Gln Phe Leu Gln Asp Phe Ala Gln Leu Pro Leu
259                               895                               900                               905
261 gat aag atg gat cca aac gtg gtc agg caa aag ttg agc aaa atg aaa      2787
262 Asp Lys Met Asp Pro Asn Val Val Arg Gln Lys Leu Ser Lys Met Lys
263                               910                               915                               920
265 acc gac ctg gag agg gat gca gtt gac tct cac tgg ctt cag caa ttc      2835
266 Thr Asp Leu Glu Arg Asp Ala Val Asp Ser His Trp Leu Gln Gln Phe
267                               925                               930                               935
269 ttt taattcttca gattagaact atcttctatt ctgtgaagct tgggggggaa      2888
270 Phe
273 tgatacttat gggttttgtg gatataactt agcctatctg taaactttca tttaaactcct      2948
275 taccccaac atgattctct gtaatcaggg gacttttgta tgcattctgt gttaatagta      3008
277 agcggtatct tatatgggtca aaaaa      3033
280 <210> SEQ ID NO: 2
281 <211> LENGTH: 939
282 <212> TYPE: PRT
283 <213> ORGANISM: Nicotiana tabacum
285 <400> SEQUENCE: 2
287 Met Asn Glu Asn Leu Glu Glu Gln Ser Lys Leu Pro Glu Leu Lys Leu
288 1                               5                               10                               15
291 Asp Ala Lys Gln Ala Gln Gly Phe Leu Ser Phe Phe Lys Thr Leu Pro
292                               20                               25                               30
295 Lys Asp Pro Arg Ala Val Arg Leu Phe Asp Arg Arg Asp Tyr Tyr Thr
296                               35                               40                               45
299 Ser His Gly Asp Asp Ala Thr Phe Ile Ala Glu Thr Tyr Tyr His Thr
300                               50                               55                               60
303 Thr Thr Ala Leu Arg Gln Leu Gly Asn Arg Ala Asp Ala Leu Ser Ser
304 65                               70                               75                               80
307 Val Ser Val Ser Arg Asn Met Phe Glu Thr Ile Ala Arg Asp Ile Leu
308                               85                               90                               95
311 Leu Glu Arg Met Asp Arg Thr Leu Glu Leu Tyr Glu Gly Ser Gly Ser
312                               100                              105                              110
315 Asn Trp Arg Leu Val Lys Ser Gly Thr Pro Gly Asn Leu Gly Ser Phe
316                               115                              120                              125
319 Glu Asp Ile Leu Phe Ala Asn Asn Glu Met Gln Asn Ser Pro Val Ile
320                               130                              135                              140
323 Ala Ala Leu Ala Pro Asn Phe Gly Gln Asn Gly Cys Glu Val Gly Leu
324 145                              150                              155                              160
327 Gly Tyr Val Asp Ile Thr Lys Arg Val Leu Gly Leu Thr Glu Phe Leu
328                               165                              170                              175
331 Asp Asp Ser His Phe Thr Asn Leu Glu Ser Ala Leu Val Ala Leu Gly
332                               180                              185                              190
335 Cys Arg Glu Cys Leu Val Pro Ala Glu Thr Gly Lys Ser Ser Glu Tyr
336                               195                              200                              205
339 Arg Pro Met Phe Asp Ala Ile Ser Arg Cys Gly Val Met Val Thr Glu
340                               210                              215                              220

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: PCT/US02/40887

DATE: 01/17/2003  
TIME: 15:17:44

Input Set : A:\257618.txt  
Output Set: N:\CRF4\01172003\PU40887.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 1,2  
Seq#:6; N Pos. 1,2,141  
Seq#:7; N Pos. 1,2  
Seq#:8; N Pos. 1,2,161  
Seq#:9; N Pos. 165,166  
Seq#:10; N Pos. 1,2  
Seq#:11; N Pos. 1,2,157,158  
Seq#:12; N Pos. 1,2  
Seq#:13; N Pos. 222  
Seq#:21; N Pos. 11,12,13,14,15,16  
Seq#:29; N Pos. 3,6,9  
Seq#:30; N Pos. 9,12